Report for 2002VI3B: Seasonal to Century Scale Climate Effects on the Water Resources of the U.S. Virgin Islands

- Articles in Refereed Scientific Journals:
    - A working paper is being developed to be submitted to the Journal of Applied Meteorology.

Report Follows:

Summary Report on

Seasonal to Century Scale Climate Effects on the Water Resources of the U.S. Virgin Islands

Problem and Research Objectives

The U.S. Virgin Islands is very sensitive to precipitation fluctuations and lies in a region that is susceptible to droughts and extreme precipitation events that can cause flooding and land slides. Understanding the cyclic nature of these events will lead to better preparedness for the population and emergency managers. Ultimately, research in this area may lead to the ability to predict future precipitation events with reasonable skill.

The main purpose of this project is to understand the behavior of rainfall process of U.S. Virgin Islands and its relations with climate general circulations. Five specific targets will be accomplished:

- Determine homogeneous climate zones using air temperature, and rainfall observations.
- Identify interactions between U.S. Virgin Islands rainfall changes with global climate changes.
- Identify significant correlations between rainfall behavior and meteorological indexes.
- Identify time series models to predict rainfall process at each coop station and compare with neural network prediction skills.
- Design and train a neural network to perform monthly rainfall predictions.

Methodology

The general methodology consists of seven major tasks: (1) Data collection (2) Estimation of missing values (3) Development of time series models (4) Identification of changes on the mean of the rainfall process (5) Identification of homogenous climatic regions (6) Designing a neural network, and (7) Comparing forecasting skills between neural network and time series models. This report describes the first four tasks. This project is expected to finish by September 2003.

Principal Findings and Significance

(1) Data Collection:
The identified data set of monthly rainfall of the Virgin Islands includes 14 coop stations and the records for most of them started in 1972. Our preliminary finding shows that the oldest stations started in 1961 and these stations are: Wintberg on St. Thomas and Annaly and Fountain on St. Croix. Table 1 shows the 14 coop stations.

Table 1. Coop Stations

	Name	I.D. Number	Records
1	Wintberg, St. Thomas	24470 679450	1961 - 2001
2	Annaly, St. Croix	24424 670240	1961 - 2001
3	Fountain Valley, St. Croix	673150	1961 - 2001
4	St. Croix FAA Hamilton	670198	1961 - 2001
5	Beth Upper New Works, St. Croix	24426 670480	1972 – 2001
6	Caneel Bay, St. John	24431 671316	1972 – 2001
7	Catheringburg, St. John	671348	1972 – 2001
8	Charlotte Amalie Harbor, St. Thomas	678905	1972 – 2001
9	Christiansted Fort, St. Croix	671740	1972 – 2001
10	Cruz Bay, St. John	671980	1972 – 2001
11	East End, St. John	672551	1972 – 2001
12	East Hill, St. Croix	24444 672560	1972 – 2001
13	Estate Fort Mylner, St. Thomas	672823	1972 – 2001
14	Granard, St. Croix	24456 673677	1972 – 2001

This work focuses on studying the monthly rainfall process of the longest four coop-stations and some well-known meteorological indexes. Figure 1 shows the rainfall observation for the most complete stations. Meteorological indexes were obtained throughout the web pages of federal agencies. These indexes exhibit monthly information during the period of 1961 to 2001. The studied meteorological variables are the following: SST in the North Atlantic (5-20°N, 60-30°W), SST in the South Atlantic (0-20°S, 30°W-10°E), SST in Tropical Equatorial (10°S-10°N, 0-360°). The SST in the equatorial Pacific: el Niño 1-2 (0-10°S, 90-80°W), el Niño 3 (5°N-5°S, 150-90°W), el Niño 4, (5°N-5°S, 160°E-150°W), and el Niño 3-4 (5°N-5°S, 170-120°W). The data set also includes the North Atlantic Oscillation index, the Artic Oscillation Index, the Sahel Rainfall Index (20-8N, 20W-10E), the North Brazil Rainfall Index, Cold tong Index, and Southern Oscillation Index.

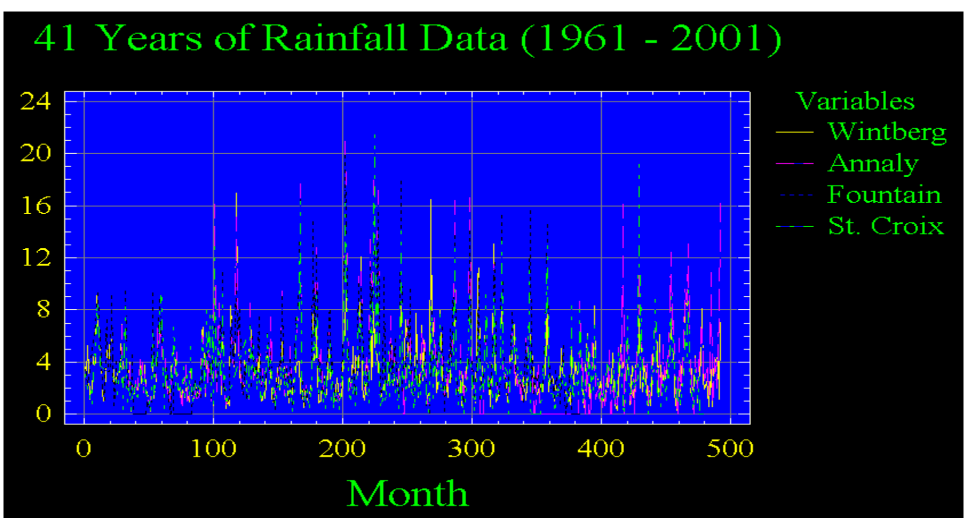

Figure 1. Monthly Rainfall Records

(2) Estimation of missing values:

The available time series exhibited some missing values. Therefore, missing observations were estimated to perform time series analysis and to develop the appropriate empirical functions. A method to perform time and spatial interpolation was implemented to estimate the missing values of the rainfall process (Ramirez et al., 2003). The interpolation algorithm is a convex combination of spatial and time interpolation methods. The convex combination can be written as follow:

$$y_{i,t} = a_i K_{i,t} + (1 - a_i) A_{i,t}\ , \qquad 0 \le a_i \le 1 \tag{1}$$

where a_i is the convex coefficient at the i^{th} station, $y_{i,t}$ is an estimate of the missing value in the i^{th} station at time t, $K_{i,t}$ is the spatial interpolation in the i^{th} station at time t obtained by using the Kriging algorithm and $A_{i,t}$ is the time interpolation in the i^{th} station at time t obtained by the seasonal autoregressive integrated and moving average (ARIMA) model (Storch and Zwiers,2001; Brockwell, and Davis, 1996; Matheron, 1979, Allard, 1998).

The convex coefficient was estimated using a cross-validation technique. A sufficient set of rainfall observations with no missing values was selected to estimate the a_i coefficients. The selected data include about 120 observations, i.e., 10 years of data. The selected time series were divided into two equal parts, the first one was used to fit the ARIMA models and the second one was used to perform time and spatial interpolation. The second part was also used to perform validation. Thus, 30% of the second part was randomly eliminated for each series and the eliminated values were declared as missing values. The Kriging algorithm and the ARIMA model were used to estimate the missing values for each station. Since the complete data set is known the a_i coefficients were estimated using linear regression techniques. Table 2 shows the estimated convex coefficient and the interpolation errors for each station. The interpolation errors are the difference between the observed rainfall value and the corresponding interpolation value.

Table 2. Convex Coefficients and Interpolation Errors

Station	Convex coefficient	Average absolute interpolation error
Wintberg	0.7523	2.7401
Annaly	0.7126	2.4516
Fountain	0.3316	2.0578
FAA, St. Croix	0.3361	2.0512

Table 2 shows that in Wintberg and Annaly stations the spatial-interpolation factor (K_{it}) is more relevant than the time-interpolation factor (A_{it}). However, the time interpolation was dominated in the remaining stations. The average absolute error of the interpolation method ranges from 2.05 to 2.74 inches. Wintberg, and St. Croix stations exhibit the maximum and the minimum interpolation errors, respectively. Since the interpolation algorithm provides a reasonable average absolute error it was used to perform the interpolation tasks to estimate all missing values in the working data series.

(3) Develop time series models:
A monthly rainfall process for a given station can be considered as a sequence of random variables. A sequence of random variables usually has deterministic and stochastic components. The deterministic components are known as seasonal and trend components. Figure 2 shows the periodogram for the rainfall of Wintberg stations. The periodogram shows that the highest spectrum is given at the frequency 0.0833 i.e., the period is 12. Periodogram also shows that at low frequency the spectrum is not significant and consequently the trend component is not significant.

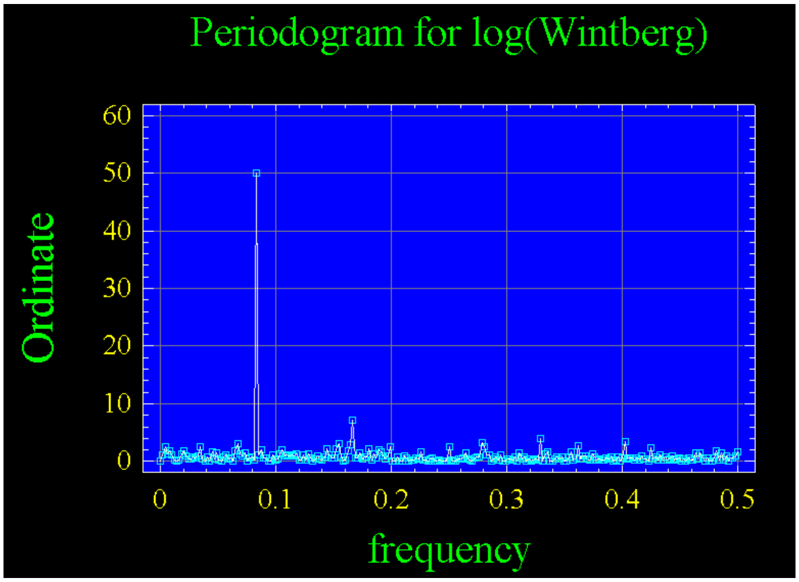

Figure 2. Periodogram for the Wintberg Station.

To model the rainfall process requires removing the seasonal component. Fitting a periodic function and then subtracting it from the original process can remove this component. A second alternative is to perform a twelve difference. Since the second alternative requires the minimum computational effort, this method was implemented. After removing the seasonal component the remaining process becomes a stationary process. However, a mathematical transformation was needed to stabilize the variance of the process. Finally the identified model includes a twelve and a first order autoregressive process. The developed model for the Wintberg station is as follows:

$$(1-B^{12})(1-\Phi B^{12})(1-fB)\,y_t = z_t$$

where y_t is the natural log of rainfall observations at month t, z_t is a white noise process with constant variance and cero mean, B is the back shift operator, $(1-B^{12})$ represents the twelve difference, Φ and f are the twelve and first autoregressive parameters which were estimated from rainfall observations. Table 3 exhibits the results of the parameter estimation and Figure 3 presents the observations versus model fitting estimation. Figure 4 shows a one-year forecast for the Wintberg station.

Table 3. Parameter Estimation (Wintberg Station)

Parameter	Estimate	Std. Error	p-value
Φ	-0.462539	0.040679	0.0000
f	0.121577	0.045714	0.0080

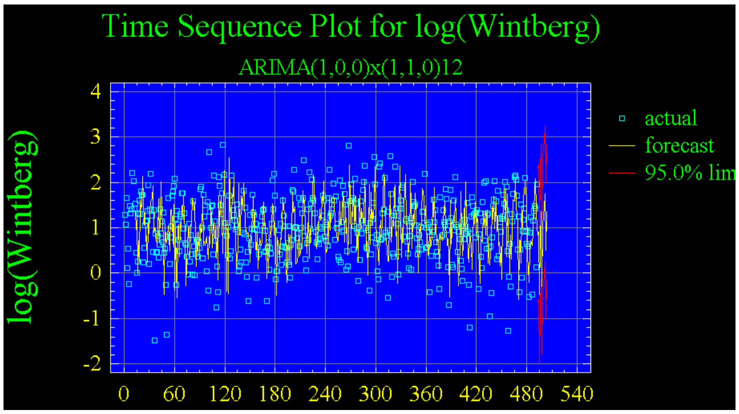

Figure 3. Rainfall Model Fitting at Wintberg Station

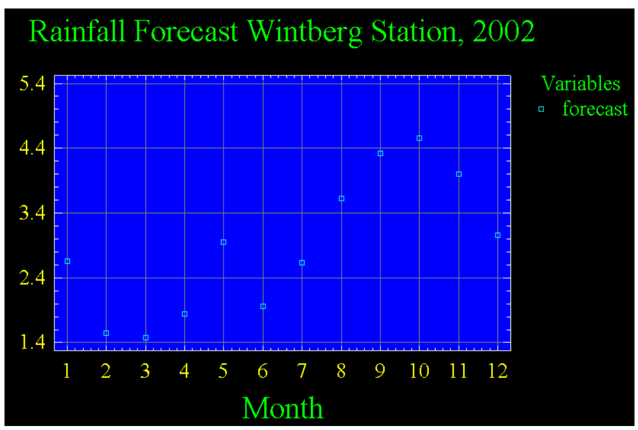

Figure 4. Rainfall Forecast for year 2002.

(4) Identify rainfall changes:

The main purpose of this task is to identify global climate changes throughout meteorological index changes and test the interaction with changes on the U.S. Virgin Islands rainfall. The available longest period of historical data is being collected for the following variables: sea surface temperature, sea level pressure, rainfall process, and air temperature. Most of the existing statistical tools to identify significant changes in the mean or in the variance of sequential information have been designed for independent time series (Cusum, Ewma, and Hotelling test). However, metereological time series are characterized by being auto- and cross-correlated processes. New statistical techniques are been tested and developed to identify significant changes on autocorrelated processes. A statistical test has been tested and compared with recent techniques (Ramirez and Sastri, 1997). Once the statistical test is selected, climate changes and Caribbean interactions will be identified. For example, it is expected to test how the North Atlantic

Oscillation index is affecting the rainfall process in the U.S. Virgin Islands; how the Sahel rainfall index interact with U.S. Virgin Island rainfall changes; how the Artic Oscillation Index interact with U.S. Virgin Islands rainfall changes. etc.

Outcomes

a. *Publications*: A working paper is being developed to be submitted to the *Journal of Applied Meteorology.*
b. *Proposal submitted/granted:* A proposal entitled: "Tropical Climate Research and Meteorological Data Management" was submitted. This proposal was submitted to the NASA program: Earth Science Research Education and Applications Solutions Network (REASON: CAN-02-OES-01). This proposal is pending. A second proposal entitled "Cloud Characterization and Time Series from Insolation Measurements Predicting" was submitted. This proposal was submitted to NASA IDEAS-ER Program and is pending. The Principal Investigator is Co-Investigator of a proposal entitled "Simulating the Hydrologic Water Balance of Puerto Rico Using a Coupled RAMS/LEAF-2/TOPMODEL/MODFLOW Modeling System." This proposal was submitted to NSF.
c. *Established partnership:* Research collaboration with NASA Goddard Space Flight Center was established. An official letter was received that research collaboration is underway. A visit to this NASA Center is planned for Summer 2003.
d. *Collaborations with internal groups:* Strong collaboration exists with remote sensing and numerical modeling groups. The statistical group is working very closely with Dr. Amos Winter, Dr. Ramon Vasquez, Dr. Jorge Gonzalez, Dr. Fernando Gilbes and Dr. Joe Eastman.
    e. *Student Participation:* Joan M. Castro is a student who is pursuing a Master of Science Degree in the Electrical Engineering Department. He is performing some statistical tests to detect climate change on meteorological indexes. Andrew S. Garcia is an undergraduate student in the Mechanical Engineering Department. He is organizing the meteorological indexes and rainfall data.

Drawbacks

The major drawback that has been identified is that the longest records started in 1961. The short records impose significant constraints to the proposed methodology.

General Future Projections

During the no-cost extension to this project, the remaining three major tasks will be completed. It is expected that a set of time series models will be developed to predict the expected rainfall one-year ahead at a specific station. The identification of significant rainfall changes and their interactions with global meteorological changes is expected. It is also expected that some correlations between meteorological indexes and rainfall processes will be identified. These correlations may suggest the physical mechanism that generates significant climate changes.

REFERENCES

Allard D., 1998. Geostatistical Classification and Class Kriging. *Journal of Geographic Information and Decision Analysis,* Vol. 2, No. 2, 77-90.

Box, G.E.P. and Jenkins, G.M. 1976. *Times Series Analysis Forecasting and Control*, Hodlen-Day Oackland, California.

Brockwell, P.J., and Davis, R.A., 1996. *Introduction to Time Series and Forecasting,*, Spring-Verlang New York Inc.

Matheron, G., 1971. *The Theory of Regionalized Variables and Applications*: Centre de Geostatistique, Fontainebleau, France.

Ramirez-Beltran, N.D., and T. Sastri. 1977. Transient Detection with an Application to a Chemical Process. *Computers & Industrial Engineering*, Vol. 32, No. 4, pp 891 -908.

Ramirez-Beltran, N.D., 2003. Time and Spatial Interpolation algorithm. A working paper to be submitted to *Apply Meteorology*.

Storch, H.V. and Zwiers, W.Z. 2001. *Statistical Analysis in Climate Research*, Cambridge University Press., Cambridge, UK.